AD_____

GRANT NUMBER DAMD17-97-1-7105

TITLE: (alpha)2(beta)1 Integrin-Induced Breast Cancer Differentiation

PRINCIPAL INVESTIGATOR: Tetsuji Kamata, M.D., Ph.D.

CONTRACTING ORGANIZATION: The Scripps Research Institute
La Jolla, California 92037

REPORT DATE: August 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1999	Annual (1 Aug 98 - 31 Jul 99)

4. TITLE AND SUBTITLE

(alpha)2(beta)1 Integrin-Induced Breast Cancer Differentiation

5. FUNDING NUMBERS

DAMD17-97-1-7105

6. AUTHOR(S)

Kamata, Tetsuji, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Scripps Research Institute
La Jolla, California 92037

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Cell-extracellular matrix (ECM) interaction is important for cell differentiation in mammary epithelial cells. Previous reports suggest that α2β1 integrin, a collagen/laminin receptor, plays a critical role in normal mammary cell differentiation as well as in the pathogenesis of breast cancer. Therefore, the modulation of α2β1-ligand interaction can be a key factor in controlling breast cancer differentiation. The purpose of this study is to clarify the role of α2β1-ligand interaction in breast cancer differentiation, and to elucidate the molecular mechanism of α2β1-ligand interaction. During the second year of this project, I found that several discontinuous amino acid residues surrounding the MIDAS (metal ion dependent adhesion site) motif in the α2 I(A) domain are critical for collagen binding. In contrast, mutation or deletion of the entire unique C-helix residues, which have previously been predicted to determine collagen binding specificity, did not affect collagen binding at all. We created a collagen/α2 I domain docking model. In this model, Asn-154, Tyr-157, Asp-219, and His-258 make direct contact with collagen. These findings will form the basis for the remainder of this study.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer, integrin, collagen, laminin	34
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

T.K. ✓ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

T.K. ✓ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

T.K. ✓ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Tetsuji Kamata 8/27/98

PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

Cell-extracellular matrix (ECM) interaction has a profound impact on cell structure and function. There is overwhelming evidence that the ECM influences mammary epithelial cell differentiation. Cell-ECM interaction is mediated through a variety of cell surface receptors including integrins. The $\alpha 2\beta 1$ integrin, a collagen/laminin receptor, is expressed in mammary epithelial cells and plays a critical role in normal mammary cell differentiation as well as in the pathogenesis of breast cancer. Modulation of $\alpha 2\beta 1$-ECM interaction could affect mammary cell differentiation. Elucidation of $\alpha 2\beta 1$-ligand binding mechanisms may lead to the development of therapeutic interventions against breast cancer.

The ligand binding site(s) in $\alpha 2\beta 1$ has been localized in the I (A) domain of the $\alpha 2$ subunit. Although several amino acid residues that constitute the metal ion-dependent adhesion site (MIDAS) are critical for ligand binding, the specific contact site for each ligand is not well characterized. I focused my effort on identifying the specific amino acid residues that make direct contact with collagen. I found that multiple discontinuous amino acid residues surrounding the MIDAS motif make direct contact with collagen. Based upon the crystal structure of the $\alpha 2$ I domain, I created a docking model of $\alpha 2$ I domain and collagen triple helical peptide. Using the same strategy, I will be able to identify residues that make contact with recombinant laminin $\alpha 1$ chain. Once the specific mutations that block either collagen or laminin are determined, I will express these mutations in a breast cancer cell line and examine their effect on differentiation. The information obtained in this study may form the basis for the development of therapeutic interventions against breast cancer.

BODY - Annual Summary

Discontinuous multiple amino acid residues in the I domain are required for collagen binding

The MIDAS face, the putative ligand binding face, of the α2 I domain is composed of four discontinuous loop structures. The βA-α1, α3-α4, and βD-α5 loops contain amino acid residues that compose the MIDAS motif, which is critical for divalent cation coordination. These amino acid residues are conserved relatively very well among different I domains. On the contrary, the amino acid sequences of the βE-α6 loop are variable. Notably, there are five amino acid insertions within this loop in collagen-binding α1, α2, and α10 integrin I domains (Fig. 1). The crystal structure of α2 I domain reveals that these inserted residues constitute an extra C-helix within the βE-α6 loop. This C-helix creates a groove on top of the MIDAS face in which a collagen triple helix is predicted to dock. Therefore the C-helix is predicted to be the major determinant of collagen binding. I reported in the previous annual report that the collagen binding sites in α2β1 are localized in the three discontinuous loops (βA-α1 loop, α3-α4 loop, and βD-α5 loop) within the I domain of the α2 subunit based on the results from α2/αL loop-swapping mutagenesis. To further characterize the collagen-contact face, I introduced multiple point mutations into the residues within the βA-α1, α3-α4, βD-α5, and βE-α6 loops. I stably expressed mutant α2 in CHO cells and examined their binding to immobilized collagen type I. Consistent with the results from loop-swapping mutagenesis, I found that Ser-153, Ile-156, and Tyr-157 in βA-α1; Gln-215 and Gly-218 in α3-α4; and Gly-255 in βD-α5, in addition to the previously reported Asp-151, Thr-221, and Asp-254, are critical for collagen binding. In addition, mutation of Asn-154, Ser-155, Asp-219, and His-258 had a mild blocking effect on collagen binding when collagen type I was immobilized at 2 μg/ml(Fig. 2). Unexpectedly, mutations in the amino acid residues in the βE-α6 loop did not affect collagen binding at all. It is possible that several amino acid residues in the βE-α6 loop have to be mutated at the same time to detect their effect on collagen binding. To rule out this possibility, I created a deletion mutant that lacks most of the residues in the βE-α6 loop, including the entire C-helix. I generated a clonal CHO cell line expressing the mutant α2 and examined its binding to plastic plates that had been coated with collagen type I at various concentrations. The del αC

mutant showed binding to collagen type I comparable with that of wild-type at all collagen immobilizing concentrations (Fig. 3). These results suggests that the del αC mutant has a collagen-binding affinity comparable to that of wild-type α2; therefore the unique C-helix does not contain an energetically important collagen contact site. Instead, the collagen contact sites are localized in the relatively conserved sequences surrounding the MIDAS motif.

α2 I domain/collagen binding model

In collaboration with Dr. Liddington, we created an α2 I domain/collagen docking model based on the current mutagenesis data. In the crystal structure, Asp-151, Ser-153, Ser-155, Thr-221, and Asp-254 are involved in the coordination of the cation. Ile-156 is totally buried in the molecule. Mutations of Gln-215, Gly-218, and Gly-255 are either buried or likely to disrupt the MIDAS motif. In contrast, Asn-154, Tyr-156, Asp-219, and His-258 are totally exposed on the surface of the molecule, suggesting that these residues make direct contact with collagen. Recently, a short synthetic triple-helical peptide, representing residues 502-516 of the collagen type I α1 chain, has been shown to support purified α2β1 and recombinant α2 I-domain.binding. The Glu and Arg residues in the GER triplet were found to be essential for recognition by the α2 I-domain. We first attached the glutamate side chain of the GER motif to the MIDAS Mg^{2+} ion, then rotated the collagen to minimize the distance between the collagen and those surface-exposed residues implicated in collagen binding (Asn-154, Tyr-157, Asp-219 and His-258) while maintaining the 2 Å bond between the glutamate oxygen and the Mg^{2+} ion and avoiding other close contact with the protein. In the current model, Asn-154, Tyr-157, Asp-219, and His-258 make direct contact with collagen, and the arginine from the GER motif makes a salt bridge to Glu-256 (Fig. 4). In addition, Tyr-285 from the C-helix also makes direct contact with collagen. However, current mutagenesis data suggest its not energetically important.

Recommendations in Relation to the Statement of Work

Thus far I have focused my effort on identifying the collagen contact site in α2β1. Using the same strategy, I plan to identify the laminin-1 contact site using a recombinant laminin α1 chain. The collagen and/or laminin-binding-defective mutants produced will be useful for elucidating the role of cell-ECM interactions in breast cancer cell differentiation.

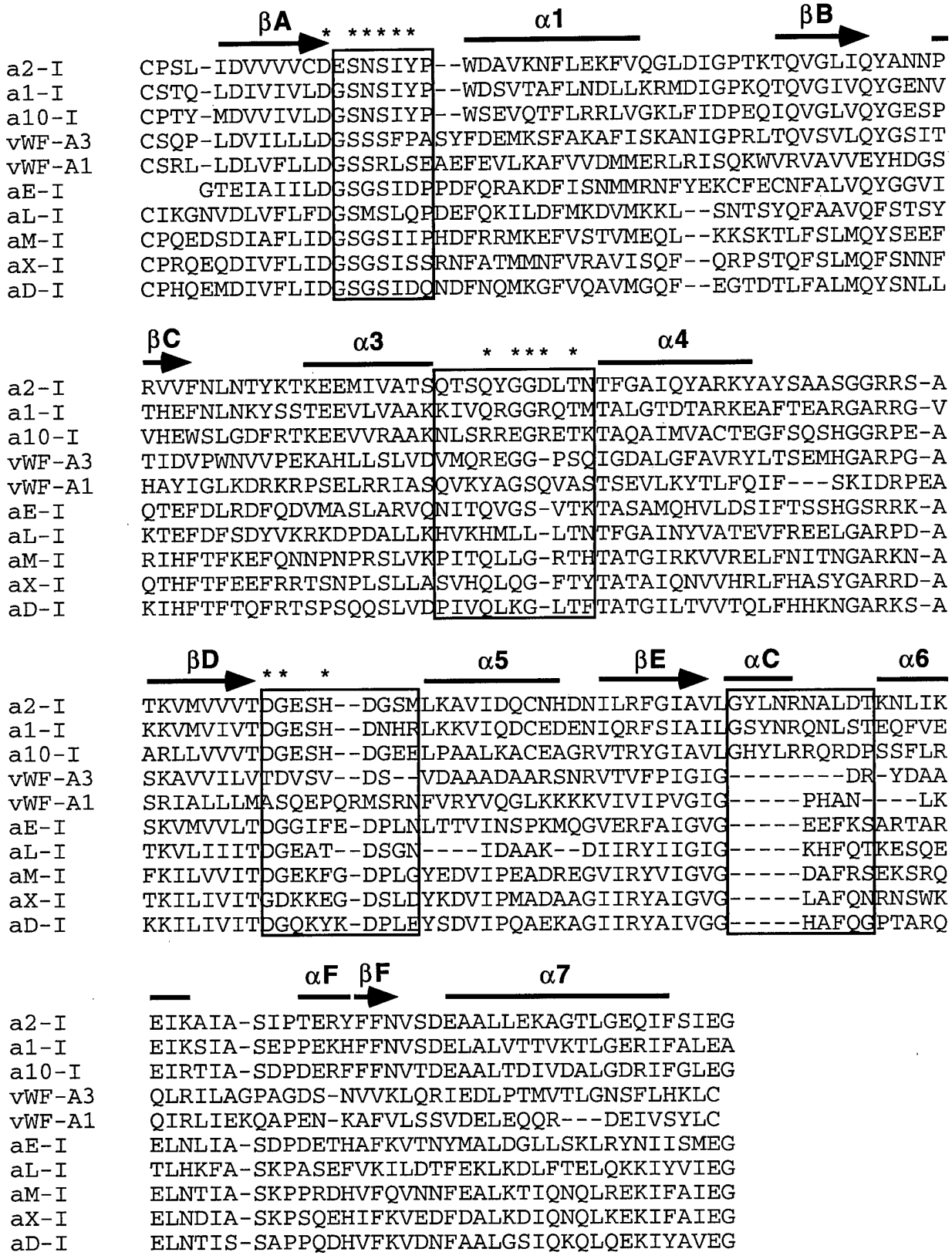


Fig. 1 **Alignment of I-domains.** The loops comprising the MIDAS motif (βA-α1, α3-α4, βD-α5) and the loop which includes αC helix (βE-α6) are outlined in boxes. The asterics indicate aminmo acid residues critical for α2β1-collagen interaction.

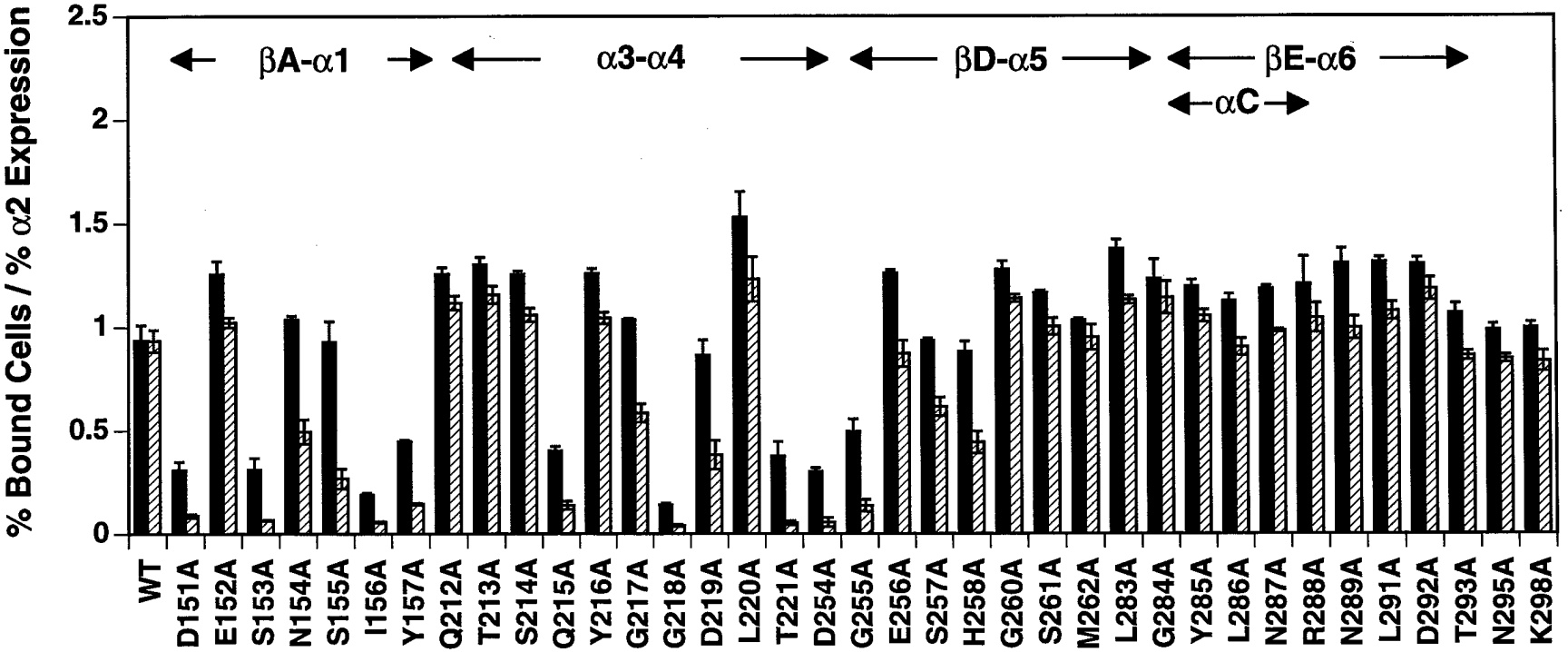


Fig. 2 **The effect of alanine mutations on the α2-expressing CHO cell adhesion to immobilized collagen type I.** Amino acid residues were substituted with Ala. CHO cells stably expressing wild-type (wt) or mutant α2 were analized for expression of α2 using FACS with HAS-4 and for adhesion to collagen type I that has been immobilized at 10 μg/ml (■) or at 2 μg/ml (▨). The data are expressed as a ratio of the persentage of collagen bound cells per the persentage of α2 positive cells.

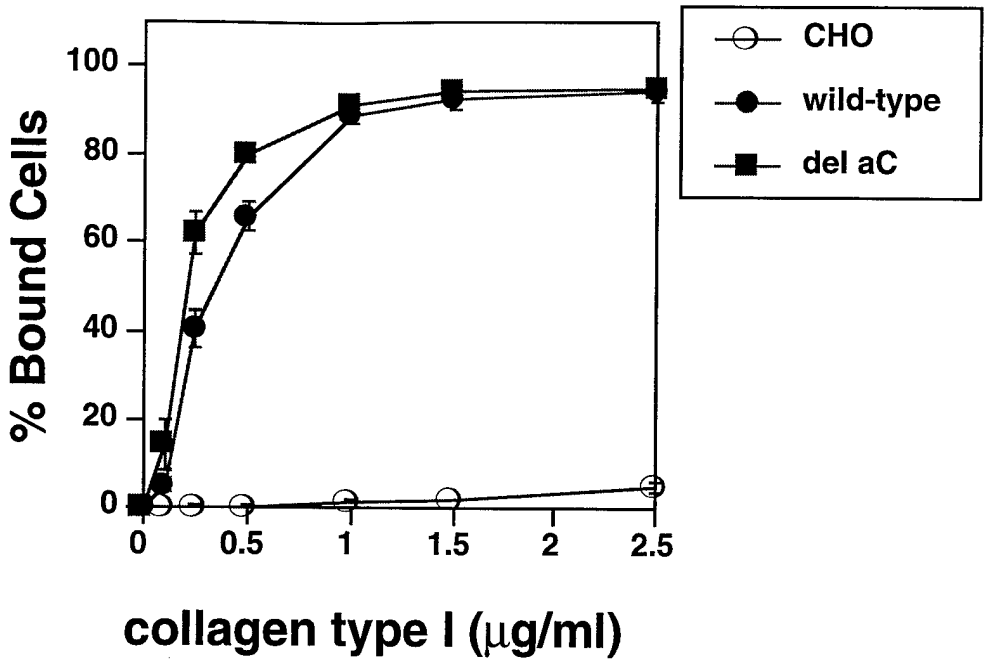


Fig. 3 **The effect of del αC mutation on α2β1-collagen type I interaction.** Cells were incubated for one hour in wells which has been precoated with different concentration of collagen type I. After washing, bound cells were quantitated by endogenous phosphatase assay.

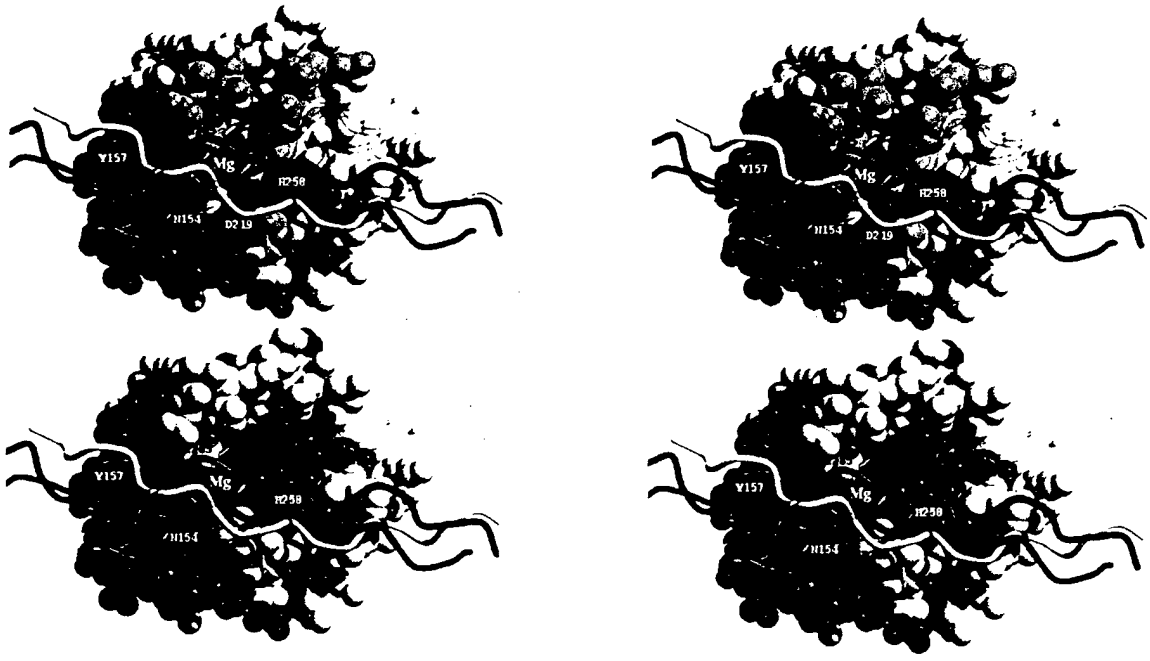


Fig. 4 **Docking model of the α2 I-domain and collagen**

All-atom representation of the α2 I domain, viewed looking down onto the MIDAS face. The Mg2+ ion is labeled at the center of the molecule. The collagen molecule is shown as a triple helical coil drawn through the Cα positions. Asn-154, Tyr-157, Asp-219, His-258 and Tyr-285 are labeled.

APPENDICES

1) Key research accomplishments
   a) Identification of the collagen contact site in the I domain of α2β1 integrin.
   b) Generation of collagen/α2 I domain docking model.

2) Reportable outcomes
   a) Manuscripts: Tetsuji Kamata, Robert C. Liddington, and Yoshikazu Takada. Interaction between collagen and the α2 I-domain of integrin α2β1: Critical role of conserved residues in the MIDAS region. *J. Biol. Chem., in press.*

3) See attached copy of the manuscript cited in 2a.

Revised M9-02432

Title: Interaction between collagen and the α2 I-domain of integrin α2β1: Critical role of conserved residues in the MIDAS region.

Authors: Tetsuji Kamata, Robert C. Liddington**, and Yoshikazu Takada*

Institution: Department of Vascular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Rd., La Jolla, CA 92037, ** The Burnham Institute, 10901 N. Torrey Pines Rd., La Jolla, CA 92037

* To whom correspondence should be sent. Yoshikazu Takada, M.D., Ph.D., Department of Vascular Biology, VB-1, The Scripps Research Institute, 10550 North Torrey Pines Rd, La Jolla, CA 92037, U.S.A.
Tel. 619-784-7122, Fax. 619-784-7323, e-mail takada@scripps.edu

This work was supported by National Institute of Health Grants GM47157 and GM49899 (to Y.T.) and by Department of the Army Grant DAMD17-97-1-7105 (to T.K.). This is a publication #12010-VB from The Scripps Research Institute.

Running title: I-domain-collagen interaction

Abbreviations used: MIDAS, metal ion dependent adhesive site; vWf, von Willebrand factor.

ABSTRACT

A docking model of the α2 I-domain and collagen has been proposed based on their crystal structures (Emsley, J., King, S., Bergelson, J. and Liddington, R. C., J. Biol. Chem., 272, 28512-28517, 1997). In this model, several amino acid residues in the I-domain make direct contact with collagen (N154, D219, L220, E256, H258, Y285, N289, L291, N295, and K298), and the protruding C-helix of α2 (residues 284-288) determines ligand specificity. Since most of the proposed critical residues are not conserved, different I-domains are predicted to bind to collagen differently. We found that deleting the entire C-helix or mutating the predicted critical residues had no effect on collagen binding to whole α2β1, with the exception that mutating N154, D219, and H258 had moderate effect. We performed further studies and found that mutating the conserved surface-exposed residues in the MIDAS (Y157 and Q215) significantly blocks collagen binding. We have revised the docking model based on the mutagenesis data. In the revised model, conserved Y157 makes contact with collagen in addition to the previously proposed N154, D219, H258, and Y285 residues. These results suggest that the collagen-binding-I-domains (e.g., α1, α2, and α10) bind to collagen in a similar fashion.

INTRODUCTION

Several integrin α chains (α1, α2, α10, αL, αM, αX, αD, and αE) have inserted I- or A- domains of about 200 amino acid residues (1-11)). Integrins α1β1 (1), α2β1 (reviewed in (12)), and α10β1 (4) have been shown to bind to collagen and/or laminin. Several function-blocking antibodies map to the I-domains of α2β1 (13) and α1β1 (14). The recombinant α2 I-domain fragment binds to collagen (15,16), and the recombinant α1 I-domain fragment binds to collagen and laminin (17). Conserved Asp and Thr residues in the α2 I-domains (D151, T221, and D254) are critical for collagen binding (15). These lines of evidence suggest that the I-domain is critically involved in collagen binding.

The crystal structures of the I-domains of the integrin αM (18), αL (19), and α2 (20) subunits, and the A1 (21,22) and A3 (23,24) domains of von Willebrand factor (vWf) have been published. This domain adopts a classic "Rossmann" fold and consists of a hydrophobic β-sheet in the middle and amphipathic α-helices on both sides. Interestingly, the integrin I-domain contains a Mg^{2+}/Mn^{2+} coordination site at its surface, which is not present in proteins with similar structures (e.g., the NAD binding domain of lactate dehydrogenase) or the vWf A1 and A3 domains (21-24). The Asp and Thr residues in α2 that have been shown to be critical for ligand binding are involved in the coordination of a divalent cation in the crystal structure (20). The α2 I-domain has a unique helix (the C-helix) protruding from the metal ion dependent adhesion site (MIDAS) that creates a groove centered on the magnesium ion. Emsley et al. proposed a model in which a collagen triple helix fits into the groove, and a Glu side chain from collagen coordinates the metal ion. In this model, the C-helix is a major determinant

for collagen binding. It was as predicted that the following I-domain residues make direct contact with collagen: N154 (the βA-α1 turn), D219 and L220 (the α3-α4 turn), E256 and H258 (the βD-α5 turn), and Y285, N289, L291, N295 and K298 (the C-helix, α6 and C-α6 turn). However, these residues are not well conserved among collagen-binding I-domains (e.g., α1, α2, and α10), suggesting that different I-domains interact with collagen in different manners. Here we show that mutation of the residues proposed to be critical for ligand binding or deletion of the entire C-helix did not significantly affect collagen binding to whole α2β1 expressed on mammalian cells except for N154, D219, and H258. In contrast, mutating several conserved MIDAS residues including Y157 significantly blocks collagen binding. We have revised the docking model based on the mutagenesis data. In the revised model, interaction between the α2 I-domain and collagen is mediated by relatively conserved residues in the MIDAS on the N-terminal side of the I-domain. Thus, it is suggested that the collagen-binding I-domains (e.g., α1, α2, and α10) bind collagen in a similar fashion.

EXPERIMENTAL PROCEDURES

Monoclonal antibodies. HAS-3 and HAS-4 (25) are generous gifts from F. Watt (Imperial Cancer Research Fund, London, UK.)

Adhesion of CHO cells to collagen.

Wells of 96 well microtiter plates (Immulon-2, Dynatech Labs., Inc., Chantlly, VA) were coated with type I collagen (2 or 10 μg/ml) at 4 °C overnight. The other protein binding sites were blocked by incubating with 1 % (w/v) bovine serum albumin (Calbiochem, CA) for 30 min at room temperature, and washing three times with PBS (10 mM phosphate, 0.15 M NaCl, pH 7.4). Cells were harvested with 3.5 mM EDTA in PBS and washed twice with Dulbecco's modified Eagle medium. 10^5 cells (in 100 μl Dulbecco's modified Eagle medium) were added to each well and incubated for 1 h at 37 °C. The wells were rinsed three times with PBS to remove unbound cells. Bound cells were quantified by assaying endogenous phosphatase activity (26).

Molecular modeling.

A model of a collagen triple helix was constructed from the crystal structure [(27), PDB entry 1cag] as previously described (20). Side-chains from the sequence of the CB3(I)5/6 peptide containing the GER motif (28) were grafted onto the collagen in standard conformations using the program TOM (29). The glutamate of one of the GER motifs was attached to the Mg^{2+} ion of the MIDAS motif via one of its carboxylate oxygens at a distance of 2.0 Å. Keeping the I-domain fixed, the collagen was then allowed to rotate around a fixed point (the glutamate oxygen) to minimize

the distance between the collagen and the side chains of those residues which showed reduced collagen binding when mutated and which were exposed on the surface of the I-domain. Unfavorably close contacts (< 2.5 Å) between the collagen and the I-domain were monitored using the program TOM. Since the triple helical nature of collagen generates three chemically distinct strands even for a homo-tripeptide (which we call the leading, middle, and trailing strands) each of these was tested separately.

Other methods.

Swapping mutagenesis was carried out using the overlap extension polymerase chain reaction (30). The positions of the α2 sequences replaced by homologous αL sequences are residues 152-157, 212-219, and 257-262 (designated βA-α1, α3-α4, and βD-α5, respectively)(Fig. 1). Deletion of residues 284-291 (designated αC del) and point mutations were created by site-directed mutagenesis using the unique site elimination method with a double stranded vector (31). The presence of mutation was confirmed by DNA sequencing. Transfection of cDNAs into CHO cells by electroporation, selection of transfected cells with G418, immunoprecipitation, and flow cytometry were carried out as previously described (32).

RESULTS AND DISCUSSION

The MIDAS of the α2 I-domain is composed of four loops (the βA-α1, α3-α4, βD-α5, and βE-α6 loops). The conserved residues D151 in the βA-α1 loop, T221 in the α3-α4 loop, and D254 in the βD-α5 loop are critical for cation coordination and ligand binding (13,15). A unique C-helix is inserted within the βE-α6 loop in the I-domains of the collagen-binding integrins α1, α2, and α10, but is not present in the I-domains of αM or αL, or the A3 domain of vWf. This C-helix has been predicted to be a major determinant for collagen binding (20).

To identify the residues in the MIDAS that are critical for collagen binding, we introduced multiple point mutations into each MIDAS loop. We also included amino acid residues (N154, D219, L220, E256, H258, Y285, N289, L291, and N295) that have been predicted to make direct contact with collagen (20). Mutant α2 was transfected into CHO cells together with a neomycin-resistant gene and selected for G-418 resistance. Cells stably expressing the mutant α2 were used for adhesion assays. Fig. 2 shows the adhesion of the mutants to collagen type I expressed as a percentage of cells adherent to collagen per percentage of human α2 positive cells (normalized adhesion to collagen). We found that mutating several residues in the βA-α1 loop (S153 and Y157) and the α3-α4 loop (Q215) blocks collagen binding. In addition, mutation of N154 and S155 in the βA-α1 loop, D219 in the α3-α4 loop, and H258 in the βD-α5 loop produced an inhibitory effect at lower collagen coating concentrations.

We also swapped the βA-α1 (residues 152-157), α3-α4 (residues 212-219), and βD-α5 loops (residues 257-262) with the corresponding sequences of αL, which does not interact with collagen (Fig. 1). These

swapping mutations did not change the conserved residues D151, T221, and D254, which are critical for cation and collagen binding. Cells expressing mutant α2 were tested for their ability to adhere to collagen. The expression of the α3-α4 swapping mutant was too low to produce reliable adhesion data (data not shown). Other mutants showed a surface-expression level comparable with that of wild type and reacted with multiple mAbs against α2 (Fig. 3a). The βA-α1 swapping mutant showed collagen binding at a background-level. Also, the βD-α5 swapping mutant showed significantly reduced collagen binding. These results are consistent with those obtained using alanine-scanning mutagenesis.

In contrast, mutation of amino acid residues in the βE-α6 loop, including the C-helix, did not have any inhibitory effect on collagen binding. Mutation of Y285, N289, L291, and N295, which are predicted to make direct contact with collagen, did not significantly affect collagen binding, even at low (2 μg/ml) collagen coating concentrations (Fig. 2). It is possible that single amino acid substitution may not be enough to induce a detectable effect on collagen binding. So, we deleted most of the βE-α6 loop, including the entire C-helix, to determine whether the C-helix is critical for ligand specificity. These mutant α2 cDNAs were stably expressed on CHO cells, and further cloned to obtain high expressors. The αC deletion mutant showed collagen binding at a level comparable to that of wild type (Fig.3a). Adhesion of the αC deletion mutant as a function of collagen coating concentration was tested. Adhesion to collagen of both wild type and αC deletion mutant α2β1 was saturated at about 1 μg/ml collagen coating concentration, indicating that the affinity to collagen is not affected by the αC deletion (Fig.3b).

These results suggest that collagen binding is mediated by relatively conserved MIDAS residues, which are located on the N-terminal side of the I-domain. S153 and S155 are involved in metal coordination, and mutating these residues would disrupt metal binding to the I-domain. Q215 is part of the MIDAS face and makes main hydrogen bond to DXSXS loop. It is likely that mutating these residues block collagen binding by disrupting metal binding to the I-domain. N154, D219, and H258 have been predicted to make direct contact with collagen in the previous model (20), although the effects of mutating these residues are moderate. Y157 is totally exposed to the surface, and this residue may make direct contact with collagen. Y157 has not been predicted to make direct contact with collagen. Mutating the other residues that are predicted to be critical for collagen binding in the proposed model (L220, E256, Y285, N289, L291, N295, and K298) has no significant effect on collagen binding. Even deletion of the entire C-helix did not significantly affect collagen binding. These results support the role of the MIDAS motif in collagen binding, since many of the mutants that affect collagen binding are predicted to disrupt the MIDAS motif. However, the present results are not fully consistent with the proposed model for collagen/α2 I-domain binding (20).

We have modified the collagen/α2 I-domain docking model based on the present mutagenesis data (Fig. 4). Recently, short synthetic triple-helical peptide, corresponding to residues 502-516 of the collagen type I α1 chain, has been shown to bind to purified α2β1 and recombinant α2 I-domain (28). The Glu and Arg residues in the GER triplet were found to be essential for recognition by α2 I-domain (28). In the current model, we first attached the glutamate side chain of the GER motif to the MIDAS Mg^{2+} ion. We then rotated the collagen to minimize the distance between the

collagen and those surface-exposed residues implicated in collagen binding (H258, Y157, D219 and N154) while maintaining the 2 Å bond between the glutamate oxygen and the Mg2+ ion and avoiding other close contacts (<2.5 Å) with the protein. It was not initially possible to make favorable hydrogen bonds with all four side chains simultaneously, so the side chains were allowed to rotate about their Cα-Cβ bonds in order to make plausible hydrogen bonds with the collagen backbone carbonyl oxygens and amide nitrogens. The collagen orientation was then refined to optimize the hydrogen bonding geometry. This procedure allowed all four I-domain side chains to make reasonable hydrogen bonds to the collagen. This model predicts that the side chain of Tyr285, which projects from the C helix into the groove, makes unavoidable contact with the collagen, and that further hydrogen bonds can be made between the Tyr hydroxyl and the collagen main chain. This revised model, which is rotated about 30 degrees from the previously published model, allows the arginine from the GER motif of the preceding strand to make a salt bridge to E256. The previous model would not allow enough space for the arginine side chain without imposing unfavorable side chain torsion angles. Since mutation of E256 does not significantly block collagen binding in the present study, this salt bridge might not be energetically important. The triple helical character of a symmetric collagen trimer generates three chemically distinct strands, which we call the leading, middle, and lagging strands. Attaching either the leading or middle strand glutamate to the Mg^{2+} ion leads to the same conclusions. Attaching the trailing strand makes a difference because there is no arginine from the preceding strand to form a salt-bridge to E256. The alternative orientation with the collagen rotated by 180 degrees is much less favorable because the arginine of the GER motif would clash sterically with the I-domain. As pointed out in our previous model, an aspartic acid

side chain in place of the GER glutamate would be too short to reach the Mg^{2+} ion without creating a large number of steric clashes. Since N154, Y157, H258 are conserved in other collagen-binding integrin I-domains (α1 and α10), it is reasonable to assume that these residues are also involved in collagen binding in these integrins (Fig. 4). This model is consistent with the observation that tyrosine and arginine are enriched in hot spots of binding energy in the protein-protein interface (33).

In the present model, the highly conserved D151 and D254 residues in the α2 I-domain are buried underneath the Mg^{2+} ion and can not contact collagen directly. We have reported that mutating these residues only partially affects collagen binding to the recombinant α2 I-domain fragment (13,15). Also, Bienkowska et al., reported that mutating the corresponding residues in the recombinant fragment of the vWf A3 domain does not affect collagen binding (23). However, the same mutation in the whole α2 molecule completely blocks collagen binding to $\alpha 2\beta 1$ (13,15). It is possible that cation coordination through these residues is critical for ligand binding in the I-domain of the integrin molecule, but not in similar domains in non-integrin structures (e.g., vWf). Consistently, the cation-binding site is not present in the vWf A3 domain (23,24). Further studies would be needed to clarify the role of the conserved Asp residues in integrin I-domains. Several other collagen-binding sites have been reported. A collagen binding surface on osteonectin has been mapped by mutagenesis: it consists of a flat surface 15 A in diameter containing polar and apolar residues (34), a crucial arginine residue, and intriguingly a Ca^{2+} binding site that might be directly involved in collagen binding. Docking and mutagenesis of Staphylococcus aureus adhesin identified another crucial arginine and a narrow groove as the binding site for collagen (35).

The collagen binding surface on vWF-A3 has not been mapped but the structurally analogous surface lacks the charged residues found in the integrin I-domains (23,24). Further studies will be required to determine whether there are any common collagen binding mechanisms.

While this paper was under review, the crystal structure of a GER-containing collagen peptide/α2 I domain complex has been solved (Emsley, J., Knight, C.G., Barnes, M.J., Farndale, R.W and Liddington, R., unpublished results). Preliminary analysis reveals that the structure is very similar to the revised model described here. Thus, the orientation and location of the collagen is as predicted, with a Glu residues from the collagen coordinating directly to the metal ion. In addition, there is an unexpected change in the C helix so that it no longer touches the collagen, in agreement with the mutagenesis results.

ACKNOWLEDGMENT

We thank F. Watt for antibodies and M. Cruz for sharing unpublished data. We also thank K.K. Tieu and W. Puzon-McLaughlin for excellent technical assistance.

REFERENCES

1. Ignatius, M. J., Large, T. H., Houde, M., Tawil, J. W., Burton, A., Esch, F., Cabonetto, S., and Reichardt, L. F. (1990) *J. Cell Biol.* **111,** 709-720

2. Briesewitz, R., Epstein, M. R., and Marcantonio, E. E. (1993) *J. Biol. Chem.* **268,** 2989-2996

3. Takada, Y., and Hemler, M. E. (1989) *J. Cell Biol.* **109,** 397-407

4. Camper, L., Hellman, U., and Lundgren-Akerlund, E. (1998) *J Biol Chem* **273**(32), 20383-9

5. Larson, R., Corbi, A. L., Berman, L., and Springer, T. A. (1989) *J. Cell Biol.* **108,** 703-712

6. Corbi, A. L., Kishimoto, T. K., Miller, L. J., and Springer, T. A. (1988) *J. Biol. Chem.* **263,** 12403-12411

7. Arnaout, M. A., Gupta, S. K., Pierce, M. W., and Tenen, D. G. (1988) *J. Cell Biol.* **106,** 2153-2158

8. Pytela, R. (1988) *EMBO J.* **7,** 1371-1378

9. Corbi, A. L., Miller, L. J., O'Connor, K., Larson, R. S., and Springer, T. A. (1987) *EMBO J.* **6,** 4023-4028

10. Van der Vieren, M., Trong, H. L., Wood, C. L., Moore, P. F., St.John, T., Staunton, D. E., and Gallatin, W. M. (1995) *Immunity* **3,** 683-690

11. Shaw, S. K., Cepek, K. L., Murphy, E. A., Russell, G. L., Brenner, M. B., and Parker, C. M. (1994) *J. Biol. Chem.* **269,** 6016-6025

12. Santoro, S. A., and Zutter, M. M. (1995) *Thromb Haemost* **74**(3), 813-21

13. Kamata, T., Puzon, W., and Takada, Y. (1994) *J. Biol. Chem.* **263,** 12403-12411 **269**(13), 9659-63

14. Kern, A., Briesewitz, R., Bank, I., and Marcantonio, E. (1994) *J. Biol. Chem.* **269,** 22811-22816

15. Kamata, T., and Takada, Y. (1994) *J. Biol. Chem.* **269,** 26006-26010

16. Tuckwell, D., Calderwood, D. A., Green, L. J., and Humphries, M. J. (1995) *J. Cell Sci.* **108,** 1629-1637

17. Calderwood, D. A., Tuckwell, D. S., Eble, J., Kuhn, K., and Humphries, M. J. (1997) *J. Biol. Chem.* **263,** 12403-12411 **272**(19), 12311-7

18. Lee, J.-O., Rieu, P., Arnaout, M. A., and Liddington, R. (1995) *Cell* **80,** 631-638

19. Qu, A., and Leahy, D. (1995) *Proc. Natl. Acad. Sci. USA.* **92,** 10277-10281

20. Emsley, J., King, S. L., Bergelson, J. M., and Liddington, R. C. (1997) *J Biol Chem* **272**(45), 28512-7

21. Emsley, J., Cruz, M., Handin, R., and Liddington, R. (1998) *J. Biol. Chem.* **263,** 12403-12411 **273**(17), 10396-401

22. Celikel, R., Varughese, K. I., Madhusudan, Yoshioka, A., Ware, J., and Ruggeri, Z. M. (1998) *Nat Struct Biol* **5**(3), 189-94

23. Bienkowska, J., Cruz, M., Atiemo, A., Handin, R., and Liddington, R. (1997) *J Biol Chem* **272**(40), 25162-7

24. Huizinga, E. G., Martijn van der Plas, R., Kroon, J., Sixma, J. J., and Gros, P. (1997) *Structure* **5**(9), 1147-56

25. Tenchini, M. L., Adams, J. C., Gilbert, C., Steel, J., Hudson, D. L., Malcovati, M., and Watt, F. M. (1993) *Cell Adhesion and Communication* **1,** 55-66

26. Prater, C. A., Plotkin, J., Jaye, D., and Frazier, W. A. (1991) *J. Cell Biol.* **112,** 1031-1040

27. Bella, J., Eaton, M., Brodsky, B., and Berman, H. M. (1994) *Science* **266**(5182), 75-81

28. Knight, C. G., Morton, L. F., Onley, D. J., Peachey, A. R., Messent, A. J., Smethurst, P. A., Tuckwell, D. S., Farndale, R. W., and Barnes, M. J. (1998) *J Biol Chem* **273**(50), 33287-94

29. Cambillau, C., Horjales, E., and Jones, T. A. (1984) *J. Mol. Graphics* **2,** 53-54

30. Horton, R. M., and Pease, L. R. (1991) in *Directed Mutagenesis; A practical approach* (McPherson, M. J., ed), IRL Press, Oxford

31. Deng, W. P., and Nickoloff, J. A. (1992) *Anal. Biochem.* **200,** 81-88

32. Takada, Y., Ylanne, J., Mandelman, D., Puzon, W., and Ginsberg, M. (1992) *J. Cell Biol.* **119,** 913-921

33. Bogan, A. A., and Thorn, K. S. (1998) *J Mol Biol* **280**(1), 1-9

34. Sasaki, T., Hohenester, E., Gohring, W., and Timpl, R. (1998) *EMBO. J* **17**(6), 1625-34

35. Symersky, J., Patti, J. M., Carson, M., House-Pompeo, K., Teale, M., Moore, D., Jin, L., Schneider, A., DeLucas, L. J., Hook, M., and Narayana, S. V. (1997) *Nat Struct Biol* **4**(10), 833-8

FIGURE LEGEND

Fig. 1: Residues/loops chosen for mutagenesis in this study.
The I-domains from integrin α1, α2, α10, αL, αM, αX, αD, and αE subunits and the vWf A1 and A3 domains were aligned. The β-strands and α-helices of the α2 I-domain are underlined. Swapped regions of α2 and αL (βA-α1: 152-157; α3-α4: 213-219; βD-α5: 257-262) are outlined in the box. The deleted region in α2 (αC del) is also outlined.

Fig. 2. Effects of point mutations on collagen binding
Cells stably expressing wt or mutant α2 were used to determine adhesion to collagen (at a 10 or 2 μg/ml coating concentration). Data are presented as % bound cells to collagen per % human α2 positive cells to normalize a2 expression. Typically 40-60% of cells are positive after selection with G-418. Previously published function-negative mutations (D151A, T221A, and D254A) are included as negative controls. These results suggest that several relatively conserved residues in the βA-α1, α3-α4, and βD-α5 loops are critical for collagen binding.

Fig. 3: Effects of swapping/deletion mutations on collagen binding.
a) Clonal CHO cells stably expressing wild type or mutant α2 were incubated in the well coated with collagen type I, or bovine serum albumin (negative control). After incubation at 37 C for 1 h, non-adherent cells were removed and bound cells were determined by assaying endogenous phosphatase. Under the conditions used more than 80% of cells adhered to fibronectin as a positive control. MFI, mean fluorescence intensity.

b) Adhesion to collagen of wild type and the αC deletion mutant α2 was determined as s function of collagen coating concentrations. The data suggest that the adhesive function of the αC deletion mutant is comparable to that of wild type.

Fig. 4.: A revised docking model of the α2 I-domain and collagen.
All-atom representation of the α2 I-domain, viewed looking down onto the MIDAS face. In the top panel, residues with mutations that reduce collagen binding are in red (surface-exposed) or pink (likely to disrupt MIDAS). Residues with mutations that have no effect on collagen binding are in cyan. The Mg^{2+} ion is shown as a gray ball. The new collagen model is shown as a colored triple helical coil (blue, green and yellow) drawn through the Cα positions. The previous model (20) is shown as a transparent triple helical coil. Certain residues referred to in the text are labeled. In the bottom panel, residues shown in red are invariant between α1, α2 and α10 I-domains.

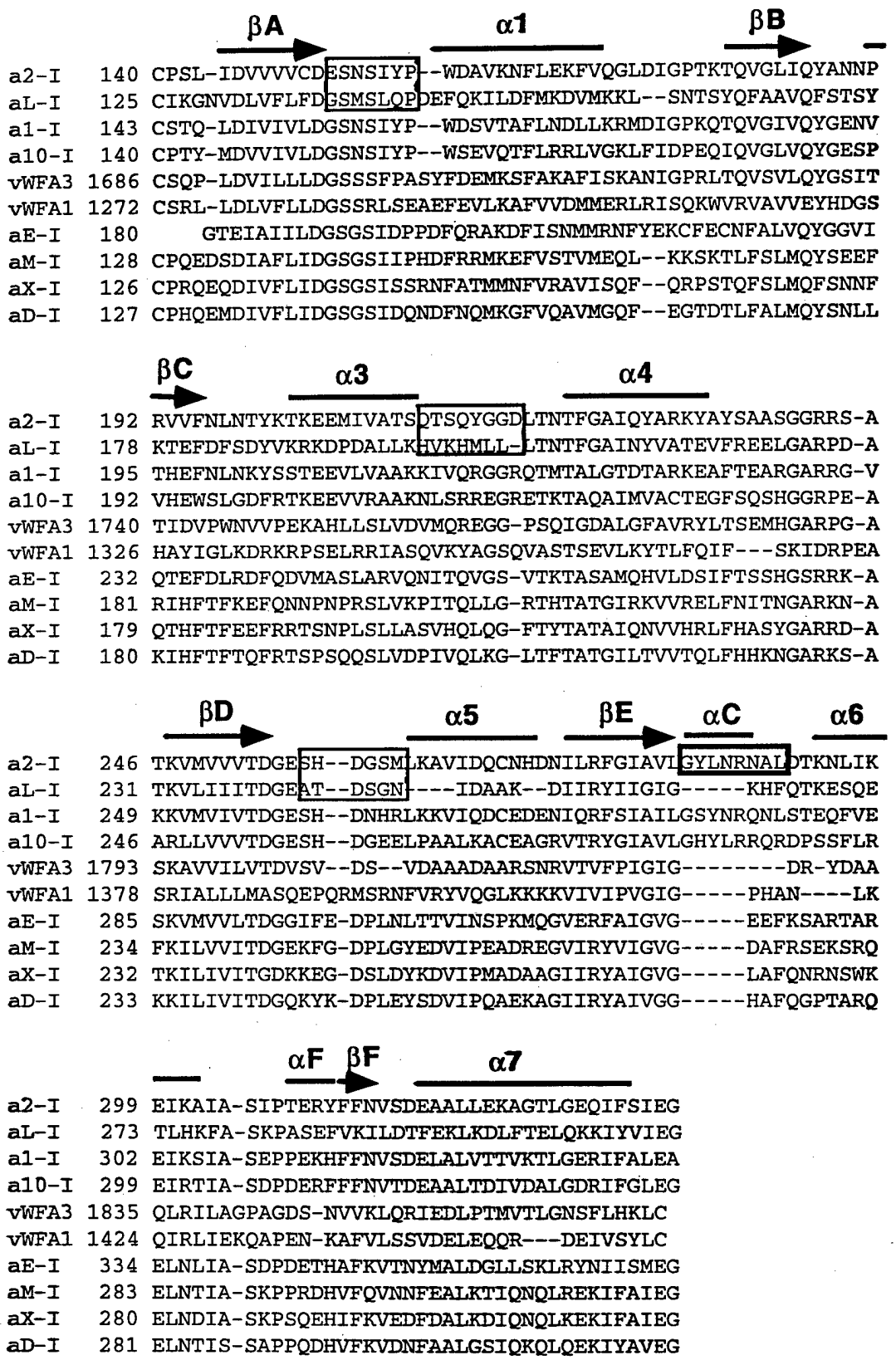
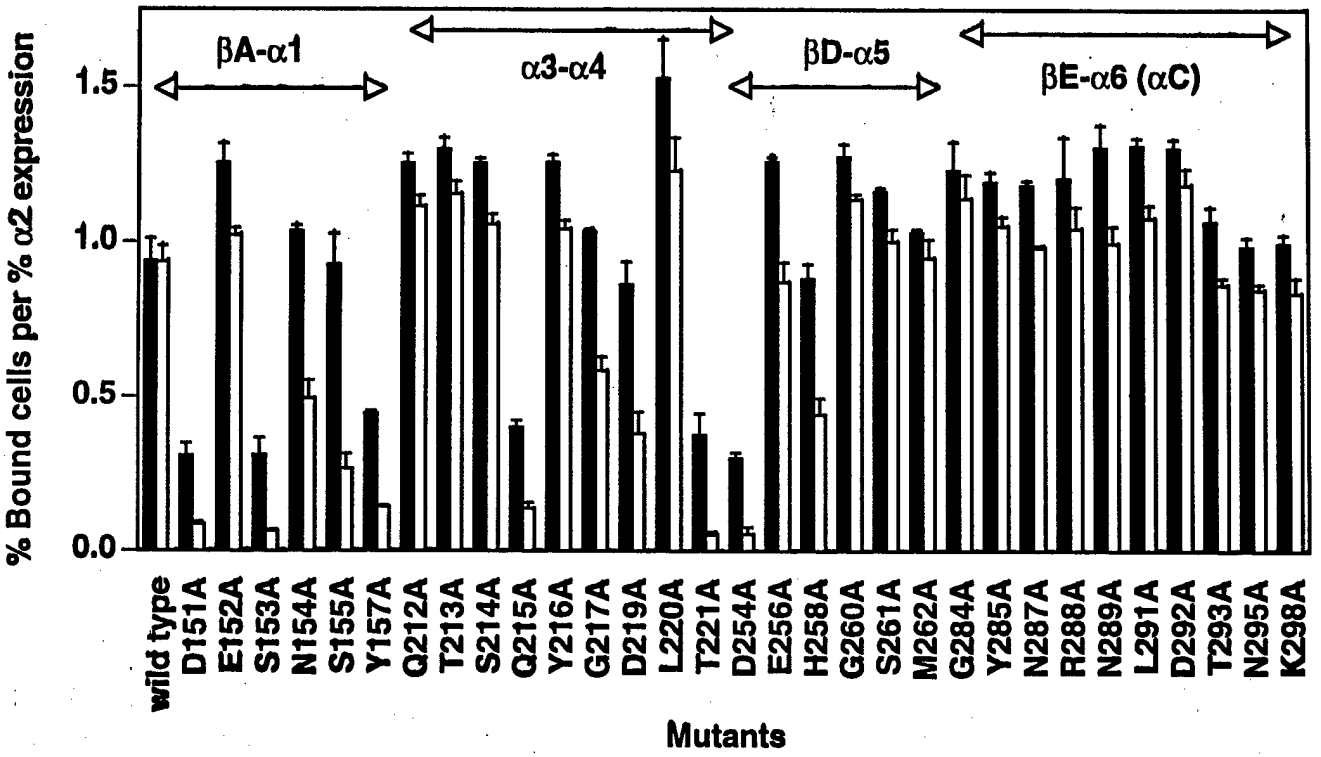
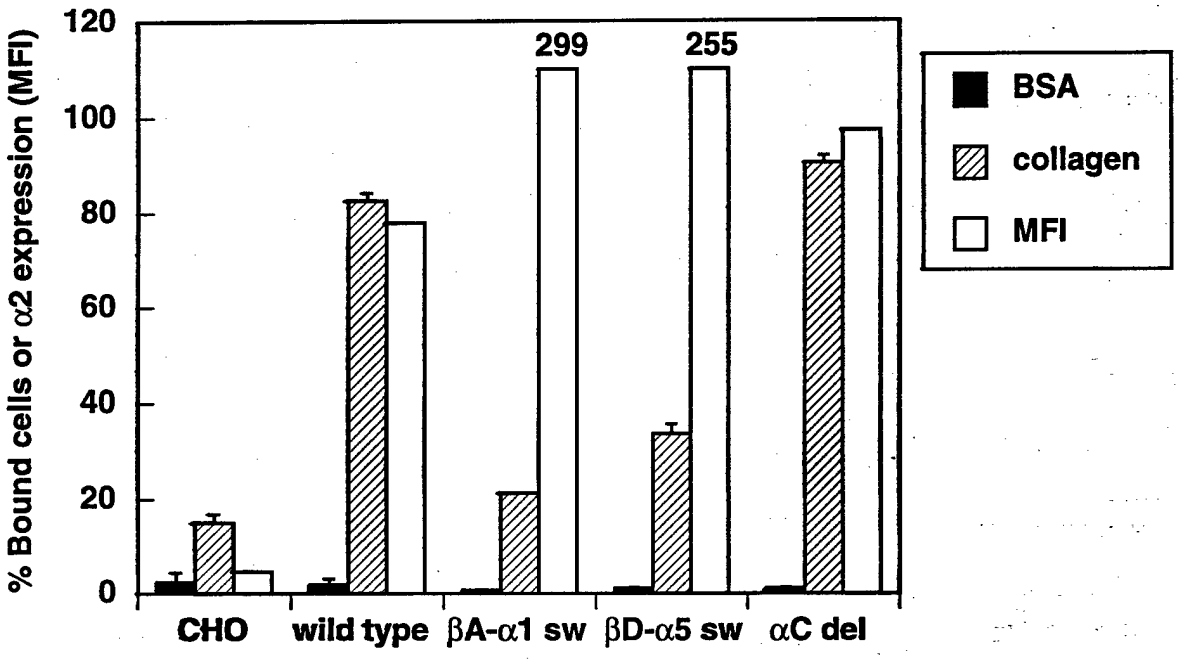

βA α1 βB
a2-I 140 CPSL-IDVVVVCDESNSIYP--WDAVKNFLEKFVQGLDIGPTKTQVGLIQYANNP
aL-I 125 CIKGNVDLVFLFDGSMSLQPDEFQKILDFMKDVMKKL--SNTSYQFAAVQFSTSY
a1-I 143 CSTQ-LDIVIVLDGSNSIYP--WDSVTAFLNDLLKRMDIGPKQTQVGIVQYGENV
a10-I 140 CPTY-MDVVIVLDGSNSIYP--WSEVQTFLRRLVGKLFIDPEQIQVGLVQYGESP
vWFA3 1686 CSQP-LDVILLLDGSSSFPASYFDEMKSFAKAFISKANIGPRLTQVSVLQYGSIT
vWFA1 1272 CSRL-LDLVFLLDGSSRLSEAEFEVLKAFVVDMMERLRISQKWVRVAVVEYHDGS
aE-I 180 GTEIAIILDGSGSIDPPDFQRAKDFISNMMRNFYEKCFECNFALVQYGGVI
aM-I 128 CPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQL--KKSKTLFSLMQYSEEF
aX-I 126 CPRQEQDIVFLIDGSGSISSRNFATMMNFVRAVISQF--QRPSTQFSLMQFSNNF
aD-I 127 CPHQEMDIVFLIDGSGSIDQNDFNQMKGFVQAVMGQF--EGTDTLFALMQYSNLL
βC α3 α4
a2-I 192 RVVFNLNTYKTKEEMIVATSQTSQYGGDLTNTFGAIQYARKYAYSAASGGRRS-A
aL-I 178 KTEFDFSDYVKRKDPDALLKHVKHMLL-LTNTFGAINYVATEVFREELGARPD-A
a1-I 195 THEFNLNKYSSTEEVLVAAKKIVQRGGRQTMTALGTDTARKEAFTEARGARRG-V
a10-I 192 VHEWSLGDFRTKEEVVRAAKNLSRREGRETKTAQAIMVACTEGFSQSHGGRPE-A
vWFA3 1740 TIDVPWNVVPEKAHLLSLVDVMQREGG-PSQIGDALGFAVRYLTSEMHGARPG-A
vWFA1 1326 HAYIGLKDRKRPSELRRIASQVKYAGSQVASTSEVLKYTLFQIF---SKIDRPEA
aE-I 232 QTEFDLRDFQDVMASLARVQNITQVGS-VTKTASAMQHVLDSIFTSSHGSRRK-A
aM-I 181 RIHFTFKEFQNNPNPRSLVKPITQLLG-RTHTATGIRKVVRELFNITNGARKN-A
aX-I 179 QTHFTFEEFRRTSNPLSLLASVHQLQG-FTYTATAIQNVVHRLFHASYGARRD-A
aD-I 180 KIHFTFTQFRTSPSQQSLVDPIVQLKG-LTFTATGILTVVTQLFHHKNGARKS-A
βD α5 βE αC α6
a2-I 246 TKVMVVVTDGESH--DGSMLKAVIDQCNHDNILRFGIAVLGYLNRNALDTKNLIK
aL-I 231 TKVLIIITDGEAT--DSGN----IDAAK--DIIRYIIGIG-----KHFQTKESQE
a1-I 249 KKVMVIVTDGESH--DNHRLKKVIQDCEDENIQRFSIAILGSYNRQNLSTEQFVE
a10-I 246 ARLLVVVTDGESH--DGEELPAALKACEAGRVTRYGIAVLGHYLRRQRDPSSFLR
vWFA3 1793 SKAVVILVTDVSV--DS--VDAAADAARSNRVTVFPIGIG--------DR-YDAA
vWFA1 1378 SRIALLLMASQEPQRMSRNFVRYVQGLKKKKVIVIPVGIG-----PHAN----LK
aE-I 285 SKVMVVLTDGGIFE-DPLNLTTVINSPKMQGVERFAIGVG-----EEFKSARTAR
aM-I 234 FKILVVITDGEKFG-DPLGYEDVIPEADREGVIRYVIGVG-----DAFRSEKSRQ
aX-I 232 TKILIVITGDKKEG-DSLDYKDVIPMADAAGIIRYAIGVG-----LAFQNRNSWK
aD-I 233 KKILIVITDGQKYK-DPLEYSDVIPQAEKAGIIRYAIVGG-----HAFQGPTARQ
αF βF α7
a2-I 299 EIKAIA-SIPTERYFFNVSDEAALLEKAGTLGEQIFSIEG
aL-I 273 TLHKFA-SKPASEFVKILDTFEKLKDLFTELQKKIYVIEG
a1-I 302 EIKSIA-SEPPEKHFFNVSDELALVTTVKTLGERIFALEA
a10-I 299 EIRTIA-SDPDERFFFNVTDEAALTDIVDALGDRIFGLEG
vWFA3 1835 QLRILAGPAGDS-NVVKLQRIEDLPTMVTLGNSFLHKLC
vWFA1 1424 QIRLIEKQAPEN-KAFVLSSVDELEQQR---DEIVSYLC
aE-I 334 ELNLIA-SDPDETHAFKVTNYMALDGLLSKLRYNIISMEG
aM-I 283 ELNTIA-SKPPRDHVFQVNNFEALKTIQNQLREKIFAIEG
aX-I 280 ELNDIA-SKPSQEHIFKVEDFDALKDIQNQLKEKIFAIEG
aD-I 281 ELNTIS-SAPPQDHVFKVDNFAALGSIQKQLQEKIYAVEG

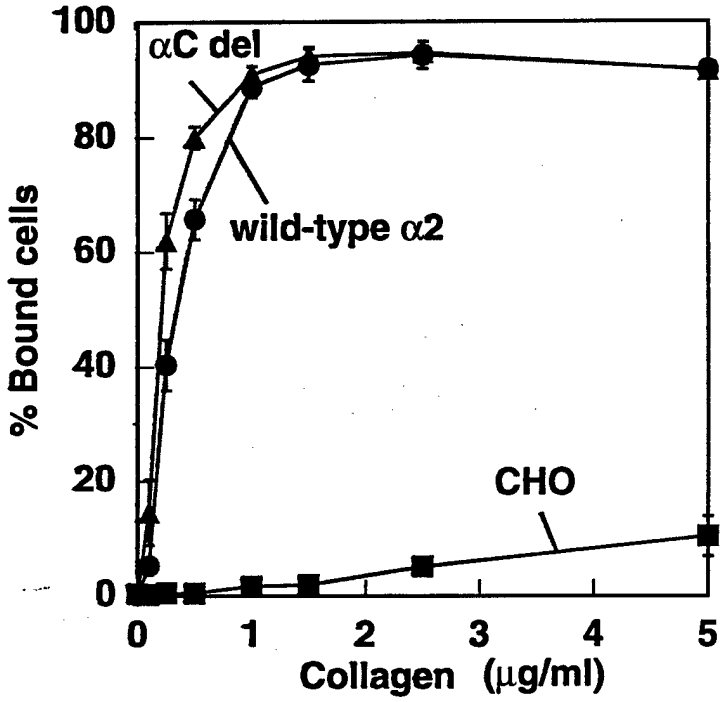
Figure 1



Kamata et al., Fig. 2



Kamata et al., Fig. 3a



Kamata et al., Fig. 3b

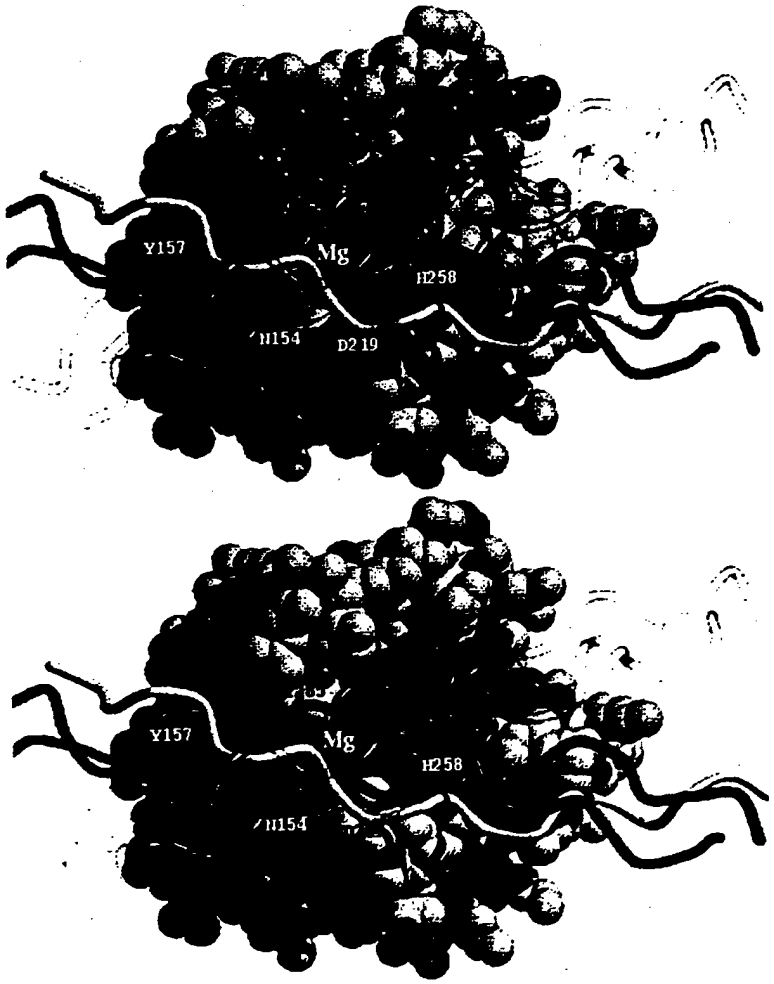


Fig. 4 Kamata et al.